REMARKS

Claims 1-5, 7-8, and 17-28 remain in this application. Claims 6, 14, 15, and 16 have been cancelled. Claims 1, 5, and 25 have been amended. Claims 26-28 have been added.

Initially, the Applicants would like to express their gratitude toward the Examiner for taking the time to discuss the present application by telephone on June 2, 2004 prior to this submission. The amendments herein reflect the subject matter discussed by telephone with the Examiner and are believed to place the claims in condition for allowance.

Before addressing the merits of the rejections based on prior art, a brief description of the present application is provided. The present application is directed toward a method and apparatus for utilizing session resources on a shared (or thin) client network environment. Computers in a network environment can be categorized as two types: servers and clients. In addition, a client can be further understood to be a thin client (in contrast with a thick client or a full-featured workstation). A thin client or a Desktop Unit (DTU) is a small, stateless, "plug and work" desktop computer whose main function is to process all input and output for the user and to manage communication with at least one server. All other client processing for the user are concentrated on a group of client servers and shared amongst a community of DTUs. The group of client servers can be called a shared client (or a consolidated client) because, although the servers are often the equivalent of larger powerful server machines, they perform the traditional role of the traditional client in a traditional client/server architecture. In addition, the shared client is shared by a large number of DTUs (that are shared by an even larger number of users on the DTUs). The removal of the traditional client processing (e.g., state maintenance and computation power) from the DTU (or thin client) into the shared client servers permits simplification of the DTU in the network

because software and hardware for performing these tasks are not needed at the DTU.

Because the DTUs are stateless (i.e., devices that process information without any knowledge of previous/subsequent information), a user's interaction with the network is managed using a persistent user session and the interaction can be instantly sent to any DTU within the network. That is, a user can be in the middle of a user session (associated with one or more services from one or more servers) on one DTU, move to another DTU, and then resume the user session exactly where the user left off. Similarly, if a DTU fails, a user can move from the failed DTU to another DTU without losing any work.

In one embodiment of the present invention, a DTU initiates a connection with any one of a plurality of available servers (e.g., a first server) in a group server environment. The DTU presents a token (which is not an IP address, but a logical identity of a user of the DTU) to the connected server. When the token is presented to the server, the server communicates with the other servers that are part of the group server environment to find one or more user sessions (within the group server environment) that are associated with the presented token. The DTU is then redirected to another server (e.g., a session server or a second server) having the one or more sessions. That is, the servers in the group server environment can communicate among each other to establish the knowledge within each server of the state about the current topology of the network and server group so that a DTU, for example, can be redirected to connect to the most recently accessed session. The state about the current topology of the network and server group is collected upon a client request (e.g., a request from the DTU using a DHCP protocol), and only those servers that are up and operating respond to the client request. In addition, all information about a token's session is created on demand and as needed by a server, so that there is no information that a server uses that is in danger of being incorrect or stale. Thus, the present invention provides a current session to a DTU hosted from a plurality of servers in a grouped server environment that are highly available and self-organizing, with no

master component, and, hence, no single point of failure.

To put it another way, the present invention is oriented around a persistent session so that the loss of any persistent user information can be avoided. That is, if a session hosting server fails, the DTU using that server switches to an alternate session server for hosting the session.

The Applicants have amended Claims 1, 5, and 25 to clarify certain features of the subject matter being claimed and/or for consistency purposes.

The Examiner rejected all of the claims (i.e., Claims 1-8 and 14-25) under 35 U.S.C. § 103(a) as being unpatentable over Narendran (U.S. Patent No. 6,070,191) in view of Andresen ("SWE: Toward a Scalable WWW Server on Multi-Computers"). These rejections are respectfully traversed.

Narendran is directed to a Round Robin DNS (or switch) 12, first and second redirection servers 14-1, 14-2, and a plurality of document servers S1, S2, S3, ..., SN. A client connection is initially made between a client needing a document from one of the document servers S1, S2, S3, ..., SN and the DNS 12. The DNS 12 then switches the client connection to one of the first and the second redirection servers 14-1, 14-2. The switch to redirection server 14-1 or 14-2 then directs the client connection to the one document server S1, S2, S3, ..., SN having the document needed by the client.

By contrast, the present invention is directed to a switch that connects a client (e.g., DTU or 800, 801, ...) to a plurality of client servers (e.g., 824, 825, 826, 827, and 828). The client servers provide client services to the client via a session. The session is hosted by one of the client servers (and associated with a token). In operation, the client 800 initially makes a client connection with a first client server (e.g., 824) for a session (or a client service). The first client server 824 then determines a second client server (e.g., 824, 825, ...) having (or hosting) the session. The first client server (e.g., 824) then redirects the client to the second client server (e.g., 824, 825, ...). Thus, one distinction between the present invention and Narendran is that the present invention provides a plurality of servers in a grouped server environment that are self-organizing,

with no master component, and, hence, no single point of failure. By contrast, Narendran still requires a master component or components, i.e., the first and second redirection servers 14-1, 14-2, to provide the redirection of the document servers.

In addition, the redirection servers 14-1, 14-2 of Narendran are only directed to redirection of the connection from the client to the document servers S1, S2, S3, ..., SN. The redirection servers 14-1, 14-2 do not provide the actual services (i.e., the document) requested by the client. The client servers of the present invention and defined claims, by contrast, not only provide the redirection services to the client but also provide the actual client services requested by the client (i.e., the DTU).

A further distinction between the present invention and Narendran is that the requested service in Narendran is the provision of a document, whereas, in the present invention, the client server provides actual client functions (e.g., state functions) that have been removed from the thin client (e.g., the DTU).

The Applicants are not sure of relevance of the Andresen reference. However, the Applicants would like to note that the Andresen reference is directed to a conventional web server. The service model of a thin client session server of the present invention and a web server are completely different. Web services and connections using a typical web server are short, transaction-oriented interactions, with no updating of state on the server. By contrast, an aim of the present invention is to provide a persistent session to a client (or DTU) that is highly available. That is, the thin client session of the present invention persists over a long time, and the session server contains the entire state of the user's session. In addition, the DTU (or thin client) of the present invention is a stateless device having no state stored in the client. Thus, unlike the service mode proposed in Andresen, a user on a client of the present invention can power-off the client, and the user can reconnect to the session of the present invention (on the same or another client and hosted by the same or another server in the failover group) and pick up exactly where the user left off, down to the position of the cursor on the screen.

The above important differences between the present invention and the cited references show that the present invention is patently distinguishable from Andresen and Narendran (whether alone or in combination). See M.P.E.P. 2143.03 ("To establish *prima facie* obviousness of a claimed invention, **all** the claim limitations must be taught or suggested by the prior art.") Moreover, in order to expedite allowance, Claim 1 has been amended to further clarify the subject matter being claimed and also to incorporate the limitations originally set forth in Claims 6 and 14-16.

More specifically, with regard to amended Claim 1, the Applicants respectfully submit that the cited references fail to suggest or disclose a method of making a computational service available in a multiple server computing environment comprising:

> exchanging information between a plurality of servers;
>
> initiating a connection between a client unit and a first server of said plurality of servers;
>
> determining **a most recently accessed session** of a plurality of sessions on said plurality of servers;
>
> determining **at said first server a location of said most recently accessed session on one of said plurality of servers**; and
>
> **redirecting said client unit via said first server to a second server of said plurality of servers having said most recently accessed session**;
>
> wherein **each** of said plurality of **sessions comprises a plurality of services requested by said client unit**;
>
> wherein **said first and second servers can each provide said plurality of services**; and
>
> wherein said plurality of services comprise **state maintenances for a user of said client unit**.

(Emphasis in bold added).

Therefore, because Narendran and Andersen fail to disclose or suggest the above defined limitations of amended independent Claim 1, the rejections of this claim, as well as Claims 2-5, 7-8, 15-25, which depend from Claim 1, should be withdrawn.

In addition, for example, Claim 22 should be allowable for the additional reason that the cited references do not disclose or suggest "maintaining said session persistently by said plurality of servers." Claim 23 is independently allowable because it recites the limitation of "wherein said client unit comprises a stateless device" and Claim 24 is independently allowable because it recites "wherein said token can identify a plurality of sessions."

New Claims 26-28 have been added. Claims 26-27 depend (directly or indirectly) on Claim 1. Claim 28 recites the limitations of a method of making a computational service available in a multiple server computing environment comprising:

> exchanging information between a plurality of servers via **a self-discovery mechanism**;
>
> **initiating a connection** between a client unit and any one of said plurality of servers **that is available to connect with said client unit by inserting a token into said client unit**;
>
> **finding a plurality of sessions associated with said token**;
>
> **determining a most recently accessed session**;
>
> **directing said client unit to a first server of said plurality of servers having said most recently accessed session**; and
>
> **redirecting said client unit via said first server to a second server** of said plurality of servers **having a next most recently accessed session if said first server fails**.

(Emphasis in bold added). It is respectfully submitted that the limitations in these new claims are neither disclosed in nor suggested by the cited references.

In view of the foregoing, the Applicants respectfully submit that Claims 1-5, 7-8, and 17-28 are in condition for allowance. Reconsideration and withdrawal of the rejections is respectfully requested, and a timely Notice of Allowability is solicited. If it would be helpful to placing this application in condition for allowance, the Applicants **strongly** encourage the Examiner to contact the undersigned counsel and conduct a telephonic interview.

To the extent necessary, Applicants petition the Commissioner for a one-month extension of time, extending to June 18, 2004, the period for response to the Office Action dated February 18, 2004. The Commissioner is authorized to charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, to Deposit Account No. 50-0639.

Respectfully submitted,

Date: June 8, 2004

Brian M. Berliner
Attorney for Applicants
Registration No. 34,549

O'MELVENY & MYERS LLP
400 South Hope Street
Los Angeles, CA 90071-2899
Telephone: (213) 430-6000